Swiss Leading House
Economics of Education • Firm Behaviour • Training Policies

Working Paper No. 91

Heterogeneous Returns to Education Over Wage Distribution: Who Profits the Most?

Simone Balestra and Uschi Backes-Gellner

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Leading House Working Paper No. 91

Heterogeneous Returns to Education Over Wage Distribution: Who Profits the Most?

Simone Balestra and Uschi Backes-Gellner

December 2013

The Swiss Leading House on Economics of Education, Firm Behavior and Training Policies is a Research Program of the Swiss State Secretariat for Education, Research, and Innovation (SERI).

www.economics-of-education.ch

Heterogeneous Returns to Education Over Wage Distribution: Who Profits the Most?*

Simone Balestra[†] and Uschi Backes-Gellner[‡]

December 2013

Abstract

This study presents evidence of heterogeneous returns to education over the wage distribution. The authors use instrumental variable quantile regression and data from the Swiss Labor Force Survey to identify the causal link between education and wages at different quantiles of the conditional distribution of wages. The results provide evidence that there is no unique causal effect of schooling and that for each individual the effect may deviate from those extensively documented by ordinary least squares or two-stage least squares. In particular, while ordinary quantile regression estimates increasing returns in the quantile index, once the endogeneity of schooling is taken into account the authors instead observe higher returns at lower quantiles of the wage distribution. Interpreting the quantile index as a measure of unobserved ability, the results suggest that higher-ability individuals have higher wages, but the slope of their wage-education profile is flatter than that for lower-ability individuals.

Keywords: Returns to education, Instrumental variable quantile regression.

JEL Classification: I21, J24, C31, C36.

---

* We are grateful to Joshua Angrist, Eric Bettinger, Giorgio Brunello, David Card, Thomas Dee, Yingying Dong, Joop Hartog, Guido Imbens, Brian Jacob, Simon Janssen, Matthias Krapf, Edward Lazear, Susanna Loeb, Hessel Oosterbeek, Aderonke Osikominu, Eric Plug, Eric Taylor, and Niels Westergaard-Nielsen for their constructive comments and valuable help. Simone Balestra benefited from a grant provided by the Swiss National Science Foundation (project PBZHP1-147219). This study is partly funded by the Swiss Federal Office for Professional Education and Technology through its Leading House on the Economics of Education, Firm Behavior and Training Policies.

† Visiting scholar at the Center for Education Policy Analysis at Stanford University and research associate at the Graduate School of Business at the University of Zurich, Switzerland. Corresponding author: simone.balestra@uzh.ch

‡ Full professor and deputy Dean at the Faculty of Economics, Business Administration and Information Technology at the University of Zurich, Switzerland.

1 Introduction

Although a positive relationship clearly exists between schooling and wages (Dickson and Harmon, 2011), the question of whether education affects individuals differently over the wage distribution is much less analyzed (Wang, 2013). Moreover, in a distributional setting, the literature has not investigated whether different types of education result in differing returns, or whether one type of education—vocational or academic—brings a return premium compared to the other at some point of the wage distribution. These questions are particularly important because a lack of information about educational tracks may lead to costly decisions for both the individual and the government (Bettinger and Baker, 2011).

To fill these gaps, in this study we first causally estimate the returns to education over the wage distribution. The analysis reveals potential heterogeneous effects of education on wages, answering the question of whether the returns are increasing, decreasing, or u-shaped across the quantiles. In a second step, we compare the returns to one extra year of academic education with the returns to one extra year of vocational education, to investigate whether one track brings a return premium at any point in the wage distribution. Such a comparison is lacking in the literature, generally because most countries do not have an extensive vocational education and training system that allows acquiring the same quality of education and the same number of years as in the academic track, or because the academic track is more prestigious or preferred than the vocational one.[1]

One notable exception is Switzerland,[2] a country with an extensive vocational education system that attracts two-thirds of the individuals in every cohort (Tuor and Backes-Gellner, 2010). The Swiss educational system allows students to achieve tertiary education degrees for both academic and vocational tracks. Therefore, using Swiss data

[1]See, e.g., Bettinger, Kremer, and Saavedra (2010) for Colombia.

[2]Other countries with similar vocational systems are Denmark and Germany (Hanushek, Woessmann, and Zhang, 2011).

allows us to shed light on heterogeneous returns to different types of education, and to answer the question of how academic and vocational education differ over the wage distribution.

The analyses that we propose address two major issues that are common for estimations of returns to education: endogeneity of education attainment (Harmon, Oosterbeek, and Walker, 2003) and heterogeneity in the returns to education (Henderson, Polachek, and Wang, 2011). While theoretical research considers both issues simultaneously (Arias, Hallock, and Sosa-Escudero, 2001; Card, 1999), empirical work often deals with only one issue at a time. To overcome the endogeneity problem, most scholars use instrumental variable estimation (Angrist and Krueger, 1991; Dickson, 2013; Harmon and Walker, 2000; Trostel, Walker, and Woolley, 2002).

However, when dealing with the heterogeneity issue, the literature has not converged to a standard method for integrating it into the analysis (Lemieux, 2008). Therefore, researchers usually rely on different methods when accounting for heterogeneity in returns to education: Sub-sample analysis (Harmon, Oosterbeek, and Walker, 2003), non-parametric estimation (Henderson, Polachek, and Wang, 2011), Bayesian hierarchical models (Koop and Tobias, 2004), and quantile regression (Fasih, Kingdon, Patrinos, Sakellariou, and Soderbom, 2012; Martins and Pereira, 2004). The first three methods focus mainly on the existence and the nature of heterogeneity, which are not the focus of this study. However, quantile regression (QR) is instead more appropriate to our research question, because QR estimates the returns to education over the wage distribution, allowing for heterogeneity through quantile-specific intercepts and quantile-specific slopes.

The use of QR in returns to education studies was hindered for many years because the endogeneity problem in QR models could not be solved. However, recent studies by Chernozhukov and Hansen (Chernozhukov and Hansen, 2008, 2013) propose an instrumental variable quantile regression (IVQR) approach that deals with both heterogeneity and endogeneity at the same time. Although the IVQR method has been applied in many research fields in economics (Atella, Pace, and Vuri, 2008; Autor, Houseman, and Kerr, 2012; Eren, 2009; Lamarche, 2011; Maynard and Qiu, 2009; Wehby, Murray, Castilla, Lopez-Camelo, and Ohsfeldt, 2009), it is relatively new to the returns to education literature. Only two studies implement IVQR to propose alternative instruments for schooling (Arabsheibani and Staneva, 2012) and to examine the inequality-reducing effect of education in China (Wang, 2013).

Exploiting a major education reform that took place in Switzerland in the 1970s, we use IVQR to causally estimate the returns to education over the wage distribution, and we compare the results with standard QR and ordinary least squares (OLS) to determine whether taking endogeneity into account changes results and conclusions. In a second step, we also distinguish between educational paths, to add a new comparison between and within academic and vocational education. In this latter comparison we are especially interested in the presence of heterogeneity, and we therefore use only conventional QR methods.[3]

The results provide evidence that there is no unique causal effect of schooling and that for each individual the effect may deviate from those extensively documented by ordinary least squares or two-stage least squares. In particular, while ordinary quantile regression estimates increasing returns in the quantile index, once the endogeneity of schooling is taken into account the authors instead observe higher returns at lower quantiles of the wage distribution. We also reveal significant heterogeneity within the academic and the vocational track, and comparing these two paths shows that academic education does not always yield higher returns. In the upper half of the wage distribution, individuals with an academic background have higher returns than individuals with a vocational background. However, at lower quantiles of the wage distribution, vocational education brings higher returns than academic education, suggesting that answering the question

[3]Nevertheless, we also performed instrumental variable (quantile) regressions, which are available upon request.

of which type of education has larger returns is not as easy as it might appear from descriptive statistics or mean regression.

The remainder of this paper proceeds as follows. Section 2 gives an overview of the theoretical background related to our research questions. Section 3 introduces the data set and presents some descriptive statistics. Section 4 shows the econometric models in detail. Section 5 presents the results, and section 6 concludes.

2 Background

In this section, we briefly present some theoretical background and empirical evidence to explain the underlying mechanisms in the individual education choice and provide a structure for our empirical analysis. We follow the theoretical model developed by Card (1999); its most interesting feature is that it considers both heterogeneity in the returns and endogeneity of education attainment in the wage equation at the same time.

Following Card, we assume that an individual chooses his level of education to maximize the following utility function defined over wage and years of education:

$$U(w, S) = \ln(w) - f(S) = \ln[g(S)] - f(S) \tag{1}$$

where $g(S)$ and $f(S)$ are increasing convex functions that represent the benefits and costs of schooling, respectively. The condition $w = g(S)$ captures the observable relationship of wage to schooling, i.e., the level of wages available at each level of education. The first order condition for optimal education is:

$$\frac{g'(S)}{g(S)} = f'(S) \tag{2}$$

In the optimum, the marginal rate of return to education equals the marginal cost. Individual heterogeneity in the optimal education choice arises from two sources: differences in the cost of education, represented by the variation in $f(S)$, and differences in the monetary benefit of education, represented by the variation in $g'(S)/g(S)$.

To characterize the well-documented fact that (log)wage is a nearly linear function of schooling that may vary across individuals,[4] we impose the following functional form to the heterogeneity components:

$$MB_i = \frac{g'(S)}{g(S)} = b_i - k_1 \cdot S_i \tag{3}$$

$$MC_i = f'(S) = r_i + k_2 \cdot S_i \tag{4}$$

where b_i and r_i are random variables with some joint distribution across the population $i = 1, 2, ..$ and k_1 and k_2 are non-negative constants. To derive an equation for the natural logarithm of wage, we integrate the expression for the marginal rate of return to education with respect to S_i:

$$\ln(w_i) = a_i + b_i \cdot S_i - \frac{1}{2} \cdot k_1 \cdot S_i^2 \tag{5}$$

where a_i is an individual-specific constant of integration.

Equation (5) is a general version of the functional form adopted in Mincer (1974). However, the salient feature of Card's model is that individual heterogeneity potentially affects both the intercept of the wage equation (through a_i) and the slope of the wage-education relation (through b_i).

This latter feature introduces three important issues into the empirical work. First, we should expect different returns to education for individuals with different levels of ability. More specifically, given that individuals acquire education up to the point where the marginal cost equals the marginal rate of return, and given that costs depend negatively

[4] Card and Krueger (1992), Heckman and Polachek (1974), and Hungerford and Solon (1987) present evidence suggesting that wages are nearly log-linear with respect to schooling. Furthermore, Park (1994) finds log-linearity to be a good approximation of the wage-schooling relationship not only at the mean but also for several quantiles of the wage distribution.

on ability, we should observe that returns to education decrease as ability increases. This means that, while higher-ability individuals have on average higher wages, the slope of their wage-education profile is flatter than that for lower-ability individuals. Second, we cannot assess the true impact of education on wages without solving the bias introduced by the endogeneity of schooling attainment, because otherwise cross-sectional estimates are (marginally) upward biased by an omitted ability variable (Heckman, Lochner, and Todd, 2006). Third, if we want to study how education affects different individuals, we need to account simultaneously for heterogeneity and endogeneity.

To incorporate these features into our analysis, we use IVQR, which estimates the causal effect of education on conditional quantiles of the wage distribution, allowing for quantile-specific intercepts and quantile-specific slopes. Given that IVQR is a relatively new method, the vast majority of the literature uses conventional QR to investigate the heterogeneous effects of education on wage (Fasih, Kingdon, Patrinos, Sakellariou, and Soderbom, 2012; Harmon, Oosterbeek, and Walker, 2003; Hartog, Pereira, and Vieira, 2001; Martins and Pereira, 2004). From these studies we conclude that returns to education vary substantially over the wage distribution, i.e., that average effects lose some important distributional features of the return to education. These studies also suggest that returns to education increase in the quantiles of wage distribution. As we can interpret the quantile index as a measure of ability (Arias, Hallock, and Sosa-Escudero, 2001; Mwabu and Schultz, 1996), this finding contrasts with what we would theoretically expect. However, the implicit assumption of exogenous schooling in conventional QR studies may explain the discrepancy between theoretical expectation and empirical findings.

The few studies applying IVQR in the return-to-education context present mixed results. Using spouse education as an instrument for education, Wang (2013) investigates the evolution of the returns in China over time, to examine the inequality-reducing effect of education. He estimates slightly decreasing returns to education over the wage distribution, ranging from 5.1 percent at the lowest quartile to 3.1 percent at the highest quartile. Proposing risky sexual behavior at an early age as a new instrument for schooling, Arabsheibani and Staneva (2012) apply IVQR to Russian data and find increasing returns over the wage distribution. Specifically, they estimate a 5 percent return at the lowest decile and a 15 percent return at the highest decile. However, when estimating the causal return to education both approaches rely on a demand-side variation in schooling, making defending the orthogonality between the instruments and the error term of the wage equation very difficult (Arcand, D'Hombres, and Gyselinck, 2005).

Pushing the analysis further, researchers and policymakers are often interested in the return to different educational paths, such as academic and vocational education. While most studies on returns to education do not consider the curriculum content of the variable years of education, policymakers—as well as students and parents—may need more information than simply the average return to a year of education, especially when they have to make decisions about different types of educational investments. In this context, a typical question is whether vocational education yields a lower or higher labor market return than an academic education of the same number of years.

Generally, the literature suggests that academic degrees have larger benefits than vocational degrees. Dearden, McIntosh, Myck, and Vignoles (2002) provide evidence on the relative value of academic and vocational qualifications in the British labor market. Their results show that the wage premium associated with academic qualifications is on average higher than that associated with vocational qualifications at the same level. Similarly, Saniter (2012) examines the returns to education for different educational groups in Germany. He finds that the return to education is 8.5 percent for the entire sample, 2.3 percent for graduates from the basic school track (vocationally oriented), and 11 percent for graduates from a higher school track (academically oriented). Focusing on non-monetary benefits of educational tracks, Hanushek, Woessmann, and Zhang (2011) find that gains in youth employment from vocational education are offset by less adaptability and consequent diminished employment later in life. Thus, over the life-cycle, academic education appears to have larger non-monetary benefits than vocational education.

However, none of these studies analyze the return to one extra year of academic education with the return to one extra year of vocational education to investigate whether one track brings a return premium, nor do they explore the possibility of heterogeneous effects between and within educational paths. Instead, they focus on qualifications and non-monetary benefits, probably because many countries do not have an education system that allows acquiring tertiary degrees in either the academic or the vocational track. In those countries, therefore, years of education are typically very different in the two tracks. As the case in Switzerland is the opposite, we complement the discussion on academic versus vocational track by revealing the heterogeneous effects of the two educational paths and by analyzing whether—and at which point of the wage distribution—one track has higher returns than the other.

3 Data and Descriptive Statistics

Before introducing the data and providing descriptive statistics, we briefly present the current Swiss education system. The education system in Switzerland consists of parallel paths divided into vocational and academic education. After nine years of compulsory schooling, about two-thirds of a youth cohort choose to pursue vocational education and training, mostly within what is called "dual system" of apprenticeship training. This kind of training generally comprises a curriculum-based on-the-job training component and a theoretical component taught at specialized vocational schools. After graduation, most of these apprentices work as skilled workers within their occupational fields. Alternatively, vocational graduates have several other options for continuing their education. They may choose to go into higher vocational education and acquire a higher vocational education degree or a university of applied sciences degree. Another post-compulsory possibility for students is to remain in the academic school system, attend academic secondary school and obtain a "Matura", a high school diploma that is a prerequisite for tertiary academic education. At tertiary academic institutions such as universities and federal institutes of technology, students can acquire degrees ranging from a bachelor's degree to a doctorate.

We base our analysis on data from the Swiss Labor Force Survey (SLFS), produced annually by the Swiss Federal Statistical Office. The data are collected by telephone interviews, and the sample is representative for the adult population permanently living in Switzerland. The main purpose of the SLFS is to provide information on employment behavior patterns and on the structure of the labor force. Strict adherence to international definitions makes Swiss data comparable with OECD, European, and U.S. data. The SLFS was conducted for the first time in 1991 and is based on a sample of about 105,000 interviews. We select the period 2000-2009, and we pool these cross-sections to build our sample.[5]

To avoid special circumstances such as those that might arise from retirement, our sample takes into account only males aged 18-60. We also restrict the sample to employed individuals to avoid misspecification resulting from people being in school or not being active in the labor force. Among the employed, to retain individuals with attachment to the labor market, we focus on fully employed workers.[6] The wage variable of the SLFS comes from the Swiss Survey on Income and Living Conditions, a very precise data source for income resulting from labor activity. Among those individuals with no missing wage, we excluded 0.5 percent of each tail end of the wage distribution to attenuate the impact of outliers and remove implausible values. Wages are expressed in Swiss Francs (CHF) throughout the entire paper, deflated to the year 2010.[7]

In the SLFS, for each individual, we can observe the entire education path from

[5]The SLFS is a rotating panel. We keep one observation per individual to prevent problems of nonrandom attrition and clustering.

[6]We use the official definition of the Swiss Federal Statistical Office, which considers an individual as fully employed if he or she has an employment of at least 90 percent.

[7]In 2010, 1 CHF = 1 USD. In Switzerland, inflation is very low and stable.

Table 1: Descriptive Statistics

Variables	Mean (1)	Std. Dev. (2)	Min (3)	Max (4)
Annual wage	81,868	41,824	12,816	390,292
Age	40.23	10.09	18.00	60.00
Years of education	13.16	2.88	7.00	21.00
Years of vocational	2.82	2.34	0.00	9.00
Years of academic	1.80	3.56	0.00	12.50
Compulsory education	0.12	0.33	0.00	1.00
Vocational education	0.65	0.48	0.00	1.00
Academic education	0.23	0.42	0.00	1.00
N	34,744			

Notes: Swiss Labor Force Survey, Authors' calculations.

compulsory education to doctorate, and we dichotomize the educational paths into academic and vocational according to the official definition of the Swiss State Secretariat for Education and Research (appendix figure A.1). After removing individuals with missing values, we are left with 34,744 observations in the sample. Tables 1 and 2 provide descriptive statistics.[8]

From the descriptive analysis on the full sample (table 1), we observe that the average worker earns an annual wage of CHF 81,868 and has acquired 13.16 years of education. In line with the statistics at the national level, in our sample 65 percent of the individuals followed a vocational path, whereas 23 percent obtained an academic degree (Tuor and Backes-Gellner, 2010). The rest of the sample (12 percent) has compulsory education as the highest educational level. Table 2 presents descriptive statistics over wage distribution, which shows the well-known positive relationship between education and wage. However, these figures do not take into account unobserved heterogeneity; in particular, differences in ability are not factored in. Therefore, descriptive results give no indication of the causal wage effects of different types of education.

[8]See appendix Table B.1 for the details on sample construction.

Table 2: DESCRIPTIVE STATISTICS OVER WAGE DISTRIBUTION

	Bottom Quartile		Second Quartile		Third Quartile		Top Quartile	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Annual wage	46,391	8,484	62,830	4,026	80,892	7,065	137,432	46,327
Age	36.79	11.05	39.14	9.94	40.93	9.22	44.06	8.53
Years of education	11.76	2.48	12.11	2.31	13.49	2.50	15.28	2.77
Years of vocational	2.31	2.04	2.87	1.97	3.47	2.27	2.66	2.84
Years of academic	0.80	2.30	0.71	2.25	1.54	3.36	4.16	4.61
Compulsory education	0.27	0.45	0.16	0.37	0.04	0.20	0.01	0.10
Vocational education	0.60	0.49	0.73	0.45	0.77	0.42	0.51	0.50
Academic education	0.13	0.34	0.11	0.31	0.19	0.39	0.48	0.50
N	8,719		8,656		8,684		8,685	

Notes: Swiss Labor Force Survey, Authors' calculations.

4 Methods

In this section, we first introduce the equations to be estimated. We use two different models: one to analyze the return to education and one to compare the academic track with the vocational track. Second, we briefly describe the estimation methods we apply, i.e., OLS, QR, and instrumental variable estimations. Given that QR and IVQR are not as common as OLS and two-stage least squares (TSLS), we give a brief overview of these two methods following Koenker and Bassett (1978) and Chernozhukov and Hansen (2013). Third, we describe and discuss the instrumental variables we use for the causal estimation of the returns to education.

4.1 The Wage Equations

To estimate the private monetary return to one additional year of education, we consider the following Mincer-like equation:

$$\ln(w_i) = \delta_0 + \beta_S \cdot S_i + \delta_1 \cdot Age_i + \delta_2 \cdot Age_i^2 + \varphi_t + u_i \tag{6}$$

In equation (6), w_i is the annual wage of individual i, S_i represents the years of education, Age_i is a proxy for labor market experience, φ_t is a set of time controls, and u_i is an error term. As is common in the literature, we exclude various determinants of wages such as tenure and industry sector, because such variables are potentially endogenous and determined by education itself (Angrist and Pischke, 2008). In model (6), the coefficient of interest is the one on the variable years of schooling β_S, which we expect to be positive and significant.

To compare the effect of one additional year of academic education to the effect of one additional year of vocational education, we develop a model similar to that used by Hartog, Pereira, and Vieira (2001) and Vandenbussche, Aghion, and Meghir (2006). Hartog et al. modify the classical Mincer wage equation and include a spline in year of education for three categories of the school system: primary, secondary, and tertiary education. With this specification, they investigate the different effects of education on wages among different levels of education. With a similar specification, Vandenbussche et al. study the effect of tertiary education on the growth rate of countries. They separate the effect of tertiary education from primary and secondary education to show that skilled labor has a higher growth-enhancing effect for countries closer to the technological frontier. In our case, we decompose the education variable as defined in model (6) into its three components: compulsory education (C), vocational education (V), and academic education (A). Thus, we can rewrite equation (6) as follows:

$$\ln(w_i) = \delta_0 + \beta_C \cdot C_i + \beta_V \cdot V_i + \beta_A \cdot A_i + \delta_1 \cdot Age_i + \delta_2 \cdot Age_i^2 + \varphi_t + u_i \tag{7}$$

In model (7), the parameters of interest are β_V and β_A. With this second specification, we compare the return premium of one additional year of vocational education with the premium of one additional year of academic education.[9] While expecting both parameters to be significant and positive is reasonable, building expectations about the comparison between the two is not straightforward, for the following two reasons. First, previous literature on the topic is scarce. Existing studies either compare higher tracks with lower tracks (Saniter, 2012) or focus on non-monetary returns (Hanushek, Woessmann, and Zhang, 2011) and returns to qualifications (Dearden, McIntosh, Myck, and Vignoles, 2002). Second, inserting this topic into a distributional framework creates an additional challenge, because—as in the case of returns to education in general—the returns to the vocational (academic) path may be heterogeneous over the wage distribution.

[9] To test whether the two coefficients are different, we perform an F-test, whose null hypothesis is: $\hat{\beta}_V - \hat{\beta}_A = 0$.

4.2 Instrumental Variable Quantile Regression

The vast majority of applied econometrics focuses on averages, and such focus partly reflects the difficulty of producing credible average causal effects. As long as the dependent variable is binary, the mean describes the entire distribution. However, many variables such as earnings have continuous distributions, which can change in response to treatments in ways that averages do not fully reveal. QR provides a straightforward, powerful tool for modeling distributional effects, even if the underlying mechanism is complex and multidimensional (Angrist and Pischke, 2008).

To allow for heterogeneous effects of education on wages, we consider the τ^{th} conditional quantile wage function hereafter (we drop the indexes for clarity):

$$Q_{\ln(w)}[\tau|X,S] = X'\alpha(\tau) + \beta(\tau)S \tag{8}$$

where X denotes all explanatory variables other than education $(1, Age_i, Age_i^2, \varphi_t)$, $\alpha(\tau)$ is the return to X at the τ^{th} quantile, $\beta(\tau)$ is the return to education at the τ^{th} quantile, and $\tau \in (0,1) \mapsto X'\alpha(\tau) + \beta(\tau)S$ is strictly increasing in τ. In equation (8) the returns to education are a function of τ, allowing for heterogeneous effects of education on wages.

Assuming the error term in the wage equation to be independent of X and S, Koenker and Bassett (1978) propose finding the best predictor of log-wage given X and S under the asymmetric least absolute deviation loss. Doing so means estimating $\alpha(\tau)$ and $\beta(\tau)$ in equation (8) by solving the following minimization problem:

$$Q_{\ln(w)}[\tau|X,S] = \underset{\alpha(\tau),\beta(\tau)}{\arg\min} E[\rho_\tau(\ln(w) - X'\alpha(\tau) - \beta(\tau)S)] \tag{9}$$

where $\rho_\tau(u_i)$ is the“check function” defined as $\rho_\tau(u_i) = [\tau - \mathbb{1}(u_i \leq 0)]u_i$. In practice, the minimization problem is solved via linear programming and implemented in many statistical packages. As previously discussed, assuming independence between the education variable and the error term may be too stringent because of potential unobserved wage determinants (i.e., ability bias). To account for potential dependence between S and u in a distributional framework, we apply the IVQR method developed by Chernozhukov and Hansen (Chernozhukov and Hansen, 2006, 2008, 2013).

As in the case of TSLS, the identification of the IVQR approach relies on the existence of a vector Z of instrumental variables that are statistically related to S but independent of the error term u. Additionally, we have to assume that, given the information (X, S), the distribution of the structural error does not vary across the endogenous state S ("rank similarity").[10] The structural error is responsible for heterogeneity of potential outcomes among individuals with the same observed characteristics. This error term determines the relative ranking of observationally equivalent individuals in the distribution of potential outcomes conditional on the individual's observed characteristics. Rank similarity differs from exact rank invariance by allowing deviations in the individual rank away from some common level. In such formulation, we assume that an individual selects an education level without knowing the exact potential outcomes. Unfortunately, we cannot test rank similarity, but this assumption is consistent with many empirical situations where the exact latent outcomes are not known before a certain treatment.

Chernozhukov and Hansen show that assuming rank similarity implies the following moment condition:

$$\mathbb{P}[\ln(w) \leq Q_{\ln(w)}(\tau|X, S)|X, Z] = \tau \tag{10}$$

and thus, in our case:

$$\mathbb{P}[\ln(w) - X'\alpha(\tau) - \beta(\tau)S \leq 0|X, Z] = \tau \tag{11}$$

The moment condition given in (11) provides a statistical restriction for use in estimating the parameters $\alpha(\tau)$ and $\beta(\tau)$. Pointing out that equation (11) is equivalent to the state-

[10] $u|X, Z \sim U(0, 1)$, i.e., for each S and S' given (X, S): $U_S \sim U_{S'}$.

ment that zero is the τ^{th} quantile of the random variable $\ln(w) - Q_{\ln(w)}(\tau|X,S)$ conditional on (X,Z), Chernozhukov and Hansen formulate the problem as finding $[\alpha(\tau),\beta(\tau)]$ so that zero is the solution to the standard quantile regression of $[\ln(w) - X'\alpha(\tau) - \beta(\tau)S]$ on (X,Z):

$$0 = \underset{f \in F}{\arg\min} E[\rho_\tau(\ln(w) - X'\alpha(\tau) - \beta(\tau)S - f(X,Z))] \tag{12}$$

where F is the class of measurable functions of (X,Z). In our empirical application, we restrict F to the values of Z_i, i.e., $f(X,Z) = Z'\hat{\gamma}$. To obtain an estimate for $\beta(\tau)$, we look for a value $\hat{\beta}$ that makes the estimated coefficient on the instrumental variable $\hat{\gamma}(\beta,\tau)$ in equation (12) as close to zero as possible using a series of conventional quantile regression.

In practice, the IVQR estimator consists of a two-step procedure: For a given value of $\beta^j(\tau)$, we first run the ordinary QR of $\ln(w) - \beta^j(\tau)S$ on X and Z to obtain the estimates $[\hat{\alpha}(\beta^j(\tau),\tau), \hat{\gamma}(\beta^j(\tau),\tau)]$. Second, we test $\hat{\gamma}(\beta^j(\tau),\tau) = 0$ and save the corresponding F-statistic, F_j. We then repeat these two steps for all the values in a pre-specified support for $\beta^j(\tau)$ and the value that minimizes the F-statistic is the IVQR estimator $\hat{\beta}(\tau)^{IVQR}$. Once we have $\hat{\beta}(\tau)^{IVQR}$, we retrieve the correspondent $\hat{\alpha}(\tau)$.[11]

The IVQR approach allows for an interpretation of the $\hat{\beta}(\tau)^{IVQR}$ as actual effects on individuals having fixed their level of unobserved heterogeneity at a given quantile. Therefore, the effect is not identified only for the set of individuals whose treatment is altered by switching the instrument from zero to one, as in the case of the IV quantile treatment estimator proposed by Abadie, Angrist, and Imbens (2002). Furthermore, the IVQR method puts no restriction of the form of the endogenous variables and instruments (i.e., they can be binary, discrete, or continuous).

[11] To obtain the point estimates and standard errors, we use both the Stata command `ivqreg` and the Matlab function `invqr`, with almost no difference between the two approaches. The codes are publicly available at http://faculty.chicagobooth.edu/christian.hansen/research/

4.3 Identification Strategy

Given the widely acknowledged endogeneity of educational attainment in the wage equation, finding valid instruments to control for this phenomenon is crucial. However, choosing suitable instruments remains a topic of great debate in the literature on returns to education (Arcand, D'Hombres, and Gyselinck, 2005; Dickson, 2013; Heckman, Lochner, and Todd, 2006). In general, an ideal instrument should be correlated with educational attainment but uncorrelated with the unobserved determinants of the wage.

The literature on returns to education used several instruments for education: quarter of birth (Angrist and Krueger, 1991), early smoking habits (Evans and Montgomery, 1994), presence or sex of siblings (Butcher and Case, 1994), college proximity (Card, 1994), parental education (Harmon and Walker, 2000), and spouse education (Trostel, Walker, and Woolley, 2002). Over the past decade, the literature has been investigating educational reforms as a source of exogenous variation in educational attainment.[12] In particular, changes in school-leaving age (Dickson, 2013; Harmon and Walker, 1999) and compulsory education expansions (Brunello, Fort, and Weber, 2009; Brunello, Fabbri, and Fort, 2013; Fang, Eggleston, Rizzo, Rozelle, and Zeckhauser, 2012) have been attracting research interest. Following this last strand of the literature, we exploit a major reform in the Swiss educational system to build our instruments and estimate the true (causal) effect of education on wages.

In Switzerland, the main responsibility for education and culture lies with the cantons, which loosely coordinate their work at the federal level. The 26 cantonal ministers of education together form a political body named the Swiss Conference of Cantonal Ministers of Education (EDK). The EDK is responsible for educational reforms, policies, and coordination at the national level. In 1970, the EDK produced an important educational reform, with the aim of standardizing certain aspects of the Swiss education from

[12] For a recent study on the impact of educational reforms on educational attainment, see Braga, Checchi, and Meschi (2013).

compulsory school through high school. This reform became official on October 29, 1970. Previously, cantons had different compulsory school duration (seven, eight, or nine years) and different school year start (either spring or fall).

The reform set nine years of compulsory education for all cantons, and mandated that the school-year start in the fall. Given that some cantons were already in line with this reform, only about half had to change their education system. Moreover, cantons did not introduce the reforms immediately after 1970. They had time to adapt their education systems in the years following the agreement, with continuous feedback to the EDK on the reform status. Thus, we are able to keep track of the introduction of the reform in each canton. Additionally, to double-check the cantonal reform status, we also contacted each canton's educational ministry. Appendix tables C.1 and C.2 give an overview of the reforms for each canton, the dates of their introduction in the canton (or not), and the ways in which the reforms modified (or not) the canton's education system.

We use the compulsory education expansion as an instrument for years of education. The empirical literature suggests that postponing the allocation of pupils to tracks yields positive effects on average educational attainment, because students stay in school longer and drop out less (Braga, Checchi, and Meschi, 2013). Similar to Brunello, Fort, and Weber (2009); Brunello, Fabbri, and Fort (2013); Fang, Eggleston, Rizzo, Rozelle, and Zeckhauser (2012), we exploit the series of natural experiments created by the staggered implementation of Switzerland's education reform as an instrument for estimating each individual's completed years of schooling. This approach obviates the problem of endogeneity due to unobservable variables that are correlated with both education and wage. Compulsory schooling instrument might not work properly for individuals at the top of the distribution, because such high wage (ability) workers may be willing to acquire more schooling independent of the education expansion. The IVQR method does not allow to compute a first stage, but we can study the reduced-form effect to find out in which parts of the wage distribution we have identification from our instrument.

Given that the effective date of the shift in school-year start also constitutes a (small) exogenous change in years of education, we might be willing to use this change as a second instrument for years of education completed. Pischke (2007) uses a similar approach for Germany, where a cohort experienced a shorter school year. In our case, however, the reform pertained all school levels from compulsory to high school, expanding the pool of "compliers." Furthermore, the individuals affected by this second reform where different from the ones affected by the compulsory schooling expansion (different cantons and/or different year of introduction). Appendix tables D.1 and D.2 present TSLS and IVQR estimates for a series of over-identified models using both instruments. However, we base our main analysis on the compulsory education expansion reform, because it is an instrument already known to the literature and because the results are qualitatively similar to the over-identified cases.

5 Results and Discussion

5.1 Causal Returns to Education Over Wage Distribution

Table 3 shows the regression outputs for model (6), which focuses on the returns to education. Mean regression (column 1, table 3) estimates a return to education of 6.7 percent, which indicates that wages rise by almost seven percent on average with each extra year of education. The effect is highly significant and not far from the few previous studies on returns to education in Switzerland, which estimate returns of about 7-8 percent (Weber and Wolter, 1999).

When we allow for heterogeneous effects of education on wage, an interesting picture emerges. QR estimates (columns 2-6, table 3) show that returns to education increase over the quantiles of the wage distribution. The return to education is 3.9 percent at the bottom decile, increasing to 6.9 percent at the median ($\tau = 0.5$), and reaching 8.9 percent

Table 3: RETURNS TO EDUCATION, OLS AND QR ESTIMATES

Variables	OLS (1)	1st Decile (2)	3rd Decile (3)	Median (4)	7th Decile (5)	9th Decile (6)
Years of education	0.067**	0.039**	0.060**	0.069**	0.076**	0.089**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.053**	0.037**	0.042**	0.048**	0.054**	0.067**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
$Age^2/100$	-0.051**	-0.039**	-0.042**	-0.046**	-0.051**	-0.064**
	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.003)
Constant	9.118**	9.474**	9.287**	9.163**	9.056**	8.801**
	(0.029)	(0.050)	(0.022)	(0.024)	(0.027)	(0.050)
Year fixed effects	YES	YES	YES	YES	YES	YES
(Pseudo) R^2	0.292	0.060	0.149	0.206	0.239	0.232
N	34,744	34,744	34,744	34,744	34,744	34,744

Notes: ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Robust standard errors are in parentheses. The dependent variable is the natural logarithm of annual wage.
Swiss Labor Force Survey, Authors' calculations.

at the top decile of the wage distribution. These results underline that average effects may hide useful information about the rest of the distribution: Further emphasizing the heterogeneous effects of education on wage, Figure 1 reports the quantile-specific returns to education from $\tau = 0.1$ to $\tau = 0.9$. Our estimated returns pattern over the wage distribution is very similar to those found by the literature for other countries (Fasih, Kingdon, Patrinos, Sakellariou, and Soderbom, 2012; Harmon, Oosterbeek, and Walker, 2003; Hartog, Pereira, and Vieira, 2001).

Table 4 presents TSLS estimates of model (6). As an instrument for years of education we use the expansion in compulsory education that took place in some cantons after 1970. The returns to education estimated by TSLS are slightly higher than OLS estimates, with a point estimate (standard error) of 9.9 percent (0.019). This result is typical in the literature on returns to education and is usually motivated by measurement error in the education variables (Card, 2001) and local average treatment effects (Imbens and Angrist, 1994; Imbens, 2010).

The coefficient on the instrumental variable in both the reduced form and first stage has the expected sign and is highly significant. In the first stage model (column 3 of

Figure 1: Returns to Education, QR Estimates

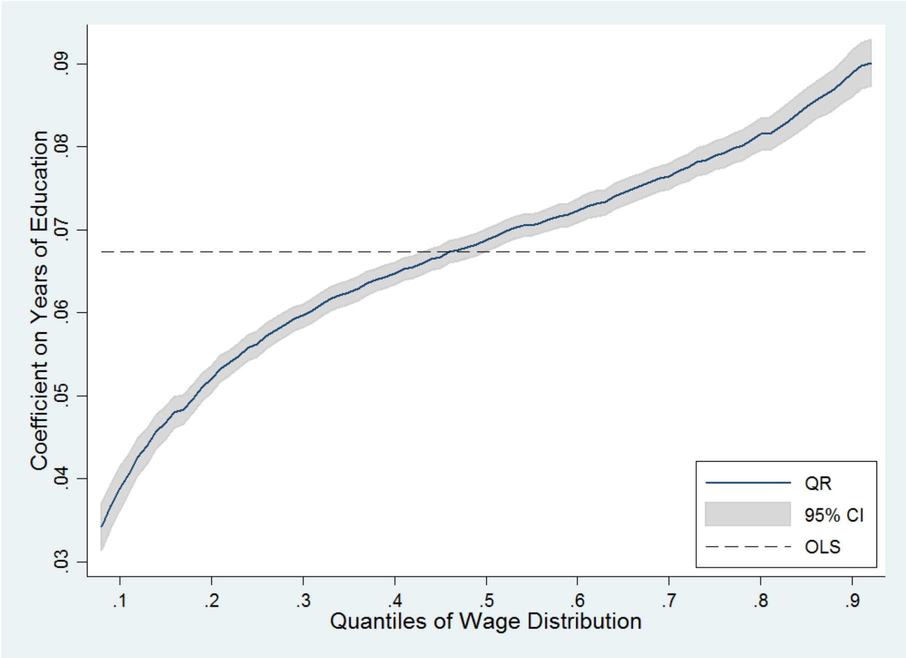


table 4), our instrument has a positive and significant effect on years of education. This finding is in line with the expectations discussed in section 4.3, and is consistent with studies that use similar instruments (Braga, Checchi, and Meschi, 2013; Brunello, Fort, and Weber, 2009; Brunello, Fabbri, and Fort, 2013; Fang, Eggleston, Rizzo, Rozelle, and Zeckhauser, 2012). In our specific case, the reforms increased educational attainment by one third of a year on average, whereas previous studies estimated an effect of about half a year. The test for excluded instruments has an F-statistic of 40.75, which is well beyond the accepted standard of 10 (Staiger and Stock, 1997). We are therefore confident about the strength of the instrumental variable. We also reject the null hypotheses of under-identification for our instrument (Kleibergen-Paap statistic).

Table 5 shows the IVQR estimates of model (6). With this regression analysis we causally estimate the impact of education on wage at a given quantile of the wage distribution. Similar to QR estimates, IVQR results also suggest that the returns to schooling vary substantially over the wage distribution. This heterogeneity is most apparent in the IVQR estimates. While both QR and IVQR approaches indicate that returns to education are heterogeneous, the shapes of the estimated returns over the quantiles are different.

Table 4: Returns to Education, TSLS Estimates

Variables	OLS (1)	Reduced Form (2)	First Stage (3)	Second Stage (4)
Years of education	0.067**			0.099**
	(0.001)			(0.019)
Age	0.053**	0.065**	0.174**	0.048**
	(0.001)	(0.002)	(0.011)	(0.003)
$Age^2/100$	-0.051**	-0.066**	-0.211**	-0.046**
	(0.002)	(0.002)	(0.014)	(0.004)
Constant	9.118**	9.748**	10.192**	8.725**
	(0.029)	(0.033)	(0.228)	(0.214)
IV–Education expansion		0.034**	0.346**	
		(0.007)	(0.054)	
Year fixed effects	YES	YES	YES	YES
R^2	0.292	0.095	0.018	0.249
N	34,744	34,744	34,744	34,744
Test for excluded instruments				
F-statistic			40.75**	
Under-identification test				
Kleibergen-Paap LM-statistic				39.90**

Notes: ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Robust standard errors are in parentheses. In columns (1), (2), and (4) the dependent variable is the natural logarithm of annual wage, in column (3) the dependent variable is years of education.

Swiss Labor Force Survey, Authors' calculations.

As in several previous studies, QR estimates exhibit increasing returns to education, indicating that returns are higher at higher quantiles of the wage distribution. However, if education is endogenous to the wage equation of model (6), conventional QR does not consistently estimate the causal effect of education on wage. IVQR estimates are instead (asymptotically) consistent under endogeneity and show that returns are decreasing over the wage distribution.

Specifically, the return to education estimated by IVQR is 18.3 percent at the first decile, decreasing to 9.6 percent at the median, and going down to an insignificant 1.6 percent at the last decile of the wage distribution. These results indicate that the largest gains to additional years of education accrue to individuals at the low end of the wage distribution. Figure 2 provides a graphical illustration of these results from $\tau = 0.1$ to $\tau = 0.9$, with a quantile interval of 0.05. The reduced-form quantile IV approach produces qualitatively similar point estimates and distributional patterns to the structural IVQR approach, indicating that our substantive results are not sensitive to the estimation procedure (Autor, Houseman, and Kerr, 2012). A look at the reduced-form effects might explain the drop in returns at the top of the wage distribution. As Figure 2 indicates, for top earners we do not have a reduced-form effect, making it impossible to compute the respective instrumental variable estimate. Therefore, the drop in return in the top decile is due to a loss of identification rather than a zero causal effect of one additional year of education. This finding is consistent with our discussion of subsection 4.3, in which we argued that our instrument would not work properly for individuals at the top of the distribution.

The IVQR estimates are also consistent with the theoretical expectations we formulated previously. As the quantile index τ can be viewed as a measure of unobserved individual ability, the IVQR results are in line with the argument that individuals acquire education up to the point where the cost equals the rate of return and where costs depend negatively on ability (Card, 1999). In this setting, we would expect the returns

Table 5: RETURNS TO EDUCATION, TSLS AND IVQR ESTIMATES

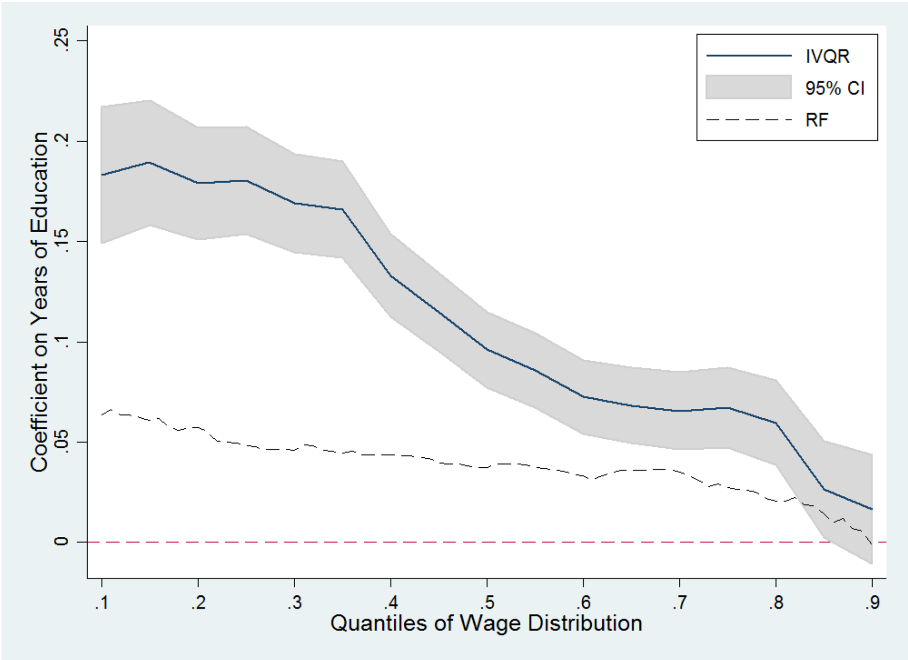
Variables	TSLS (1)	1^{st} Decile (2)	3^{rd} Decile (3)	Median (4)	7^{th} Decile (5)	9^{th} Decile (6)
Years of education	0.099**	0.183**	0.169**	0.096**	0.066**	0.016
	(0.019)	(0.017)	(0.013)	(0.010)	(0.010)	(0.044)
Age	0.048**	0.005	0.026**	0.046**	0.055**	0.100**
	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.007)
$\text{Age}^2/100$	-0.046**	0.005	-0.022**	-0.045**	-0.052**	-0.097**
	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)	(0.009)
Constant	8.725**	7.781**	8.048**	8.828**	9.178**	9.166**
	(0.199)	(0.076)	(0.055)	(0.042)	(0.043)	(0.476)
Year fixed effects	YES	YES	YES	YES	YES	YES
Reduced form effect	0.034**	0.063**	0.046**	0.037**	0.035**	-0.001
	(0.007)	(0.012)	(0.006)	(0.006)	(0.007)	(0.011)
N	34,744	34,744	34,744	34,744	34,744	34,744

Notes: ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Robust standard errors are in parentheses. The dependent variable is the natural logarithm of annual wage.
Swiss Labor Force Survey, Authors' calculations.

to education to be decreasing in ability, with the lower-ability individuals having the highest return to education—which is exactly the pattern estimated by IVQR. Moreover, interpreting the quantile index as an ability measure is also consistent with the notion that individuals with higher ability are likely to generate higher wages regardless of their educational level. Conversely, individuals with lower unobserved ability would gain more from the training provided by formal education. Our estimates suggest that higher-ability individuals indeed have higher wages, but the slope of their wage-education profile is flatter than that for lower-ability individuals.

Using a different instrument (shift in school-year start) and a combination of two instruments (compulsory education expansion and shift in school-year start) does not largely affect the estimated return (see appendix table D.1). However, with multiple instruments we have a gain in the precision of the education coefficient and we can test for over-identification. The p-value of the Hansen statistic is always not significant, indicating that we cannot reject the null hypothesis that all our model assumptions are fulfilled—including the validity of the instruments.

Figure 2: Returns to Education, IVQR Estimates



5.2 Heterogeneous Returns Between and Within Types of Education

We now focus on the comparison between educational paths. Table 6 gives an overview of the OLS and QR estimates of model (7). Column 1 of Table 6 presents OLS regressions, which estimate a return to vocational education of 6.8 percent and a return to academic education of 7.1 percent. These coefficients gather the effect of an extra year of vocational (academic) education on wage, filtering out the effect of compulsory schooling. By performing an F-test, we reject the null hypothesis of equal coefficients ($p = 0.00$), i.e., at the mean, the effect of one additional year of academic education on wage is larger than the effect on one additional year of vocational education. The question is whether modeling on average loses some important features of this comparison. Therefore, we bring the discussion into a distributional framework.

Columns 2-6 of Table 6 present the QR estimates for model (7) at various quantiles of the wage distribution. The first result is that, as in Table 3, returns to both vocational and academic education are increasing in the quantiles of the wage distribution. However, the increasing pattern and the magnitude of the estimated effects are significantly different. At the lower quantiles of the wage distribution, vocational education has a statistically significant return premium in comparison to academic education. From the fourth decile on, the situation is reversed: Academic education has higher returns for one additional year of schooling. Thus, in the upper part of the wage distribution, academic education brings a significant premium compared to vocational education.

In particular, at the bottom decile, the return to one extra year of vocational education is 5.0 percent, whereas the return to one additional year of academic education is only 4.1 percent. We reject the null hypothesis of equal coefficients at each level of significance ($p = 0.00$). At the third decile the situation is different, with an estimated return of about 6.4 percent for both academic and vocational tracks ($p = 0.66$). At the median, the returns to vocational and academic educations are 6.9 percent and 7.3 percent, respectively. Similarly to OLS, at the median we reject the null hypothesis of equal coefficients, with a p-value of 0.00. At the top decile, academic education brings a return of 9.6 percent, while vocational education has an estimated return of 8.3 percent. The difference between the estimated coefficients is statistically significant ($p = 0.00$). Figure 3 provides graphical support complementing the Table 6 results that we just discussed, comparing OLS estimates with QR estimates across the entire wage distribution, estimated for all quantiles from $\tau = 0.1$ to $\tau = 0.9$.

For a better understanding of the academic premium, we rewrite model (7) as a function of the difference between the two educational tracks, with vocational education as the reference category. While doing so prevents us from seeing the pattern of vocational and academic educations separately, the transformation allows estimating confidence intervals for the academic premium. Figure 4 plots the academic premium over the wage distribution, along with its 95 percent confidence intervals.

One potential explanation for these results is the skill formation of vocational and academic educations. Indeed, while the vocational education system provides a set of

Table 6: Returns to Vocational and Academic Education, OLS and QR

Variables	OLS (1)	1^{st} Decile (2)	3^{rd} Decile (3)	Median (4)	7^{th} Decile (5)	9^{th} Decile (6)
Compulsory education	-0.032**	-0.062**	-0.038**	-0.028**	-0.018**	-0.015*
	(0.003)	(0.006)	(0.003)	(0.003)	(0.004)	(0.007)
Vocational education	0.068**	0.050**	0.064**	0.069**	0.075**	0.083**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Academic education	0.071**	0.041**	0.064**	0.073**	0.082**	0.096**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.045**	0.029**	0.035**	0.041**	0.046**	0.058**
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
$Age^2/100$	-0.043**	-0.031**	-0.035**	-0.039**	-0.043**	-0.053**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Constant	10.133**	10.474**	10.259**	10.145**	10.043**	9.921**
	(0.045)	(0.081)	(0.042)	(0.039)	(0.046)	(0.081)
Year fixed effects	YES	YES	YES	YES	YES	YES
(Pseudo) R^2	0.309	0.072	0.163	0.220	0.250	0.243
F-statistic $\hat{\beta}_V = \hat{\beta}_A$	11.71**	23.63**	0.190	22.94**	51.06**	62.09**
N	34,744	34,744	34,744	34,744	34,744	34,744

Notes: ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Robust standard errors are in parentheses. The dependent variable is the natural logarithm of annual wage.
Swiss Labor Force Survey, Authors' calculations.

Figure 3: Returns to Vocational and Academic Education, QR Estimates

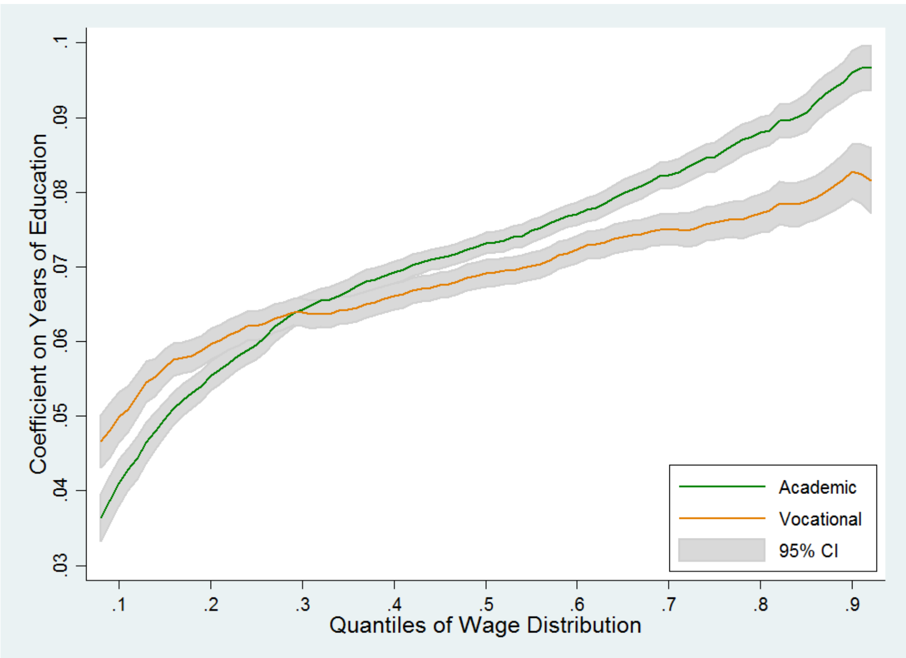
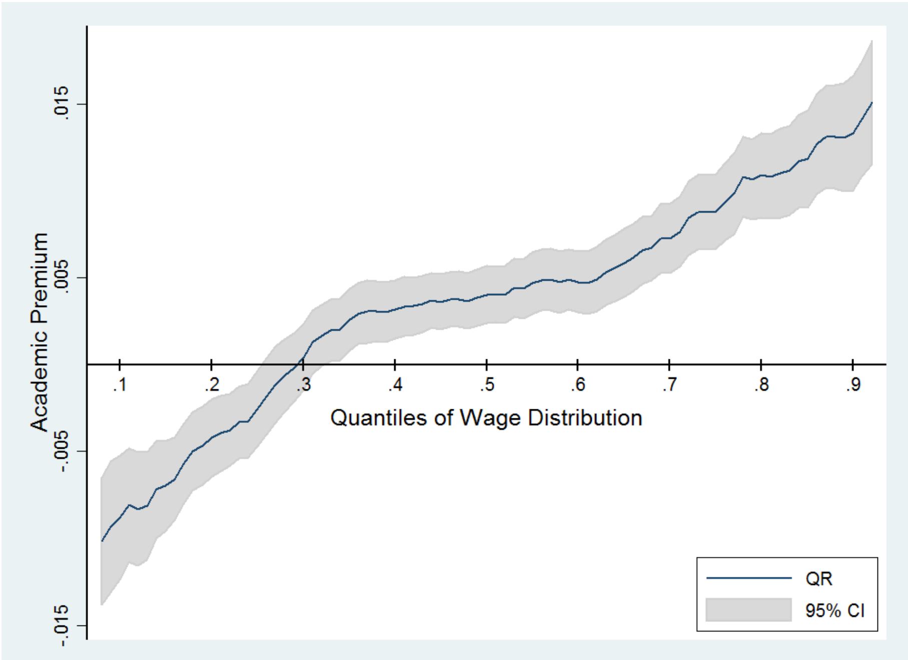


Figure 4: Academic Education Premium, QR Estimates



skills that are specific to the job that the apprentices are learning (Busemeyer and Trampusch, 2012), in academic education the exploitation of the acquired skills strongly depends on whether or not the workers are using them in the labor market (Dearden, McIntosh, Myck, and Vignoles, 2002). In addition, vocational education is likely a better fit for students at the lower part of the wage distribution, because those students learn contents that better match and complement their innate abilities (Rosenbaum and Rosenbaum, 2013). As a consequence, at the lower quantiles of the wage distribution, vocational education brings a return premium because individuals with an academic education in this part of the distribution have a relative disadvantage in the job they are performing. Conversely, at some point in the wage distribution (in our case $\tau = 0.4$) academic education, as opposed to vocational education, starts generating a return premium because workers have the capacity of fully exploiting their skills in the labor market.

Given that we are more interested in the presence of heterogeneity and because we did not find appropriate instrumental variables for both academic and vocational education, we do not claim that the estimated effects in the between-within path comparison are causal. We nevertheless performed some simple two-stage quantile regressions[13] that show lower returns to academic education at the bottom of the wage distribution and a return premium of the academic path in the upper part of the wage distribution. As an instrument for academic education we use a dummy that equals one if the canton of residence has a university. Dee (2004) and Card (1993) use a similar approach. For academic education we exploit regional variation in preference for vocational education compared to academic education, as in Rupietta and Backes-Gellner (2012).

6 Conclusions

This study presents evidence of heterogeneous returns to education over the wage distribution. We use instrumental variable quantile regression and data from the Swiss Labor Force Survey to isolate the causal link between education and wage at different quantiles of the conditional distribution of wages. Our results provide significant evidence that no unique causal effect of schooling exists and that for each individual the effect may be above or below the estimates extensively documented using OLS or TSLS, depending on his position in the wage distribution and his unobservable wage determinants.

In particular, while ordinary QR results indicate that returns to education are increasing in the quantile index, once we take the endogeneity of schooling into account, we instead observe higher returns at lower quantiles of the wage distribution. Interpreting the quantile index as a measure of unobserved ability, our findings suggest that less able individuals profit more form one additional year of education. While higher-ability individuals have on average higher wages, the slope of their wage-education profile is flatter than that for lower-ability individuals. This finding indicates, as discussed by Ashenfelter and Rouse (1998), that more able individuals acquire more schooling because they face

[13]Results are available upon request. We do not report them because the instruments we use are only *arguably* exogenous. We used the two-stage QR approach of Chen and Portnoy (1996), based on Powell (1983) early work.

lower marginal costs, not because they receive higher marginal benefits.

From a methodological point of view, one noteworthy result of our analysis is that a reduced-form quantile IV approach, akin to TSLS, produces qualitatively similar estimates to the structural IVQR approach, which is based on stronger assumptions. The comparability of these estimates indicates that our core results are not sensitive to the estimation procedure.

We also investigate the potential heterogeneity in the returns within and between different educational paths. Exploiting the unique feature of the Swiss educational system that allows students achieve tertiary education degrees for both academic and vocational tracks, we complement the existing literature by confirming that, at the mean, academic education brings higher returns. However, if we examine the returns over the wage distribution, we observe two relevant—and until now unknown—features of the returns to vocational and academic education.

First, we reveal significant heterogeneity within each educational path, with both vocational and academic educations presenting increasing returns over the wage distribution. Second, a comparison between the two tracks shows that academic education does not always yield higher returns. In the upper part of the wage distribution, individuals with an academic background have higher returns than individuals with a vocational background. However, at lower quantiles of the wage distribution, vocational education brings higher returns than academic education. These results imply that answering the question of whether academic education yields higher labor market returns than vocational education is not as easy as it might have once appeared from descriptive statistics or mean regression. Indeed, the answer depends on the individual's position in the conditional wage distribution.

Our work can be extended in a number of ways. First, analyzing the evolution over time of the quantile returns to education, and what impact the returns have on the structure of wages, would be valuable. According to our results, education should have an inequality-reducing effect over time, because individuals with lower ability (i.e., those at the lower quantiles of the wage distribution) appear to profit more from formal education. However, such inquiry is complicated by the likelihood that the endogeneity and measurement error biases change over time.

Second, in line with several cross-country studies conducted for example by Martins and Pereira (2004) and by Trostel, Walker, and Woolley (2002), researchers and policy-makers might use an international comparison to study how the causal returns to education change with different wage distributions and education systems. Third, researchers could explore the potential non-linear relationship between education and wages by allowing the returns to differ not only between educational paths but also across education levels, as, for example, in Buchinsky (1994); Hartog, Pereira, and Vieira (2001).

A fourth, and compelling, extension to our work would be evaluating the impact of changes in the distribution of education on quantiles of the unconditional (marginal) distribution of wages. Doing so would help estimate the effect of one additional year of schooling on the entire wage distribution, not only at a given quantile. However, to shed light on this topic, we would need an adaptation of the unconditional QR approach (Firpo, Fortin, and Lemieux, 2009) to instrumental variables estimation—an adaptation not yet available.

This study shows that typical estimates of the mean return to education provide a relatively incomplete characterization of the impact of education on labor market outcomes and thus constitute a weak guide for public policy. Similarly, distributional analyses using ordinary QR also constitute an inappropriate tool for describing the true impact of education on wages, because they do not control for unobserved heterogeneity. Our results suggest that the net impact of education on the long-run distribution of income does not necessarily depend on the initial distribution of ability across the population, and we empirically support the argument that formal education partially compensates for differences in innate abilities and early life conditions.

References

Abadie, A., J. Angrist, and G. Imbens (2002): "Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings," *Econometrica*, 70(1), 91–117.

Angrist, J. D., and A. B. Krueger (1991): "Does compulsory school attendance affect schooling and earnings?," *The Quarterly Journal of Economics*, 106(4), 979–1014.

Angrist, J. D., and J.-S. Pischke (2008): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arabsheibani, R., and A. V. Staneva (2012): "Returns to education in Russia: Where there is risky sexual behaviour here is also an instrument," IZA Discussion Paper 6726, Institute for the Study of Labor (IZA).

Arcand, J.-L., B. D'Hombres, and P. Gyselinck (2005): "Instrument choice and the returns to education: New evidence from Vietnam," Labor and Demography 0510011, EconWPA.

Arias, O., K. F. Hallock, and W. Sosa-Escudero (2001): "Individual heterogeneity in the returns to schooling: Instrumental variables quantile regression using twins data," *Empirical Economics*, 26(1), 7–40.

Ashenfelter, O., and C. Rouse (1998): "Income, schooling, and ability: Evidence from a new sample of identical twins," *The Quarterly Journal of Economics*, 113(1), 253–284.

Atella, V., N. Pace, and D. Vuri (2008): "Are employers discriminating with respect to weight? European evidence using quantile regression," *Economics and Human Biology*, 6(3), 305–329.

Autor, D. H., S. N. Houseman, and S. P. Kerr (2012): "The effect of work first job placements on the distribution of earnings: An instrumental variable quantile regression approach," NBER Working Paper 17972, National Bureau of Economic Research.

Bettinger, E., and R. Baker (2011): "The effects of student coaching in college: An evaluation of a randomized experiment in student mentoring," NBER Working Paper 16881, National Bureau of Economic Research.

Bettinger, E., M. Kremer, and J. E. Saavedra (2010): "Are educational vouchers only redistributive?," *The Economic Journal*, 120(546), F204–F228.

Braga, M., D. Checchi, and E. Meschi (2013): "Educational policies in a long-run perspective," *Economic Policy*, 28(73), 45–100.

Brunello, G., D. Fabbri, and M. Fort (2013): "The causal effect of education on body mass: Evidence from Europe," *Journal of Labor Economics*, 31(1), 195–223.

Brunello, G., M. Fort, and G. Weber (2009): "Changes in compulsory schooling, education and the distribution of wages in Europe," *The Economic Journal*, 119(536), 516–539.

Buchinsky, M. (1994): "Changes in the US wage structure 1963-1987: Application of quantile regression," *Econometrica*, 62(2), 405–58.

Busemeyer, M. R., and C. Trampusch (2012): *The Political Economy of Collective Skill Formation*. Oxford University Press.

Butcher, K. F., and A. Case (1994): "The effect of sibling sex composition on women's education and earnings," *The Quarterly Journal of Economics*, 109(3), 531–563.

Card, D. (1993): "Using geographic variation in college proximity to estimate the return to schooling," NBER Working Paper 4483, National Bureau of Economic Research.

——— (1994): "Earnings, schooling, and ability revisited," NBER Working Paper 4832, National Bureau of Economic Research.

——— (1999): "The causal effect of education on earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 30, pp. 1801–1863. Elsevier.

——— (2001): "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 69(5), 1127–1160.

Card, D., and A. B. Krueger (1992): "Does school quality matter? Returns to education and the characteristics of public schools in the United States," *The Journal of Political Economy*, 100(1), 1–40.

Chen, L.-A., and S. Portnoy (1996): "Two-stage regression quantiles and two-stage trimmed least squares estimators for structural equation models," *Communications in Statistics–Theory and Methods*, 25(5), 1005–1032.

Chernozhukov, V., and C. Hansen (2006): "Instrumental quantile regression inference for structural and treatment effect models," *Journal of Econometrics*, 132(2), 491–525.

——— (2008): "Instrumental variable quantile regression: A robust inference approach," *Journal of Econometrics*, 142(1), 379–398.

——— (2013): "Quantile models with endogeneity," *Annual Review of Economics*, 5(1), 57–81.

Dearden, L., S. McIntosh, M. Myck, and A. Vignoles (2002): "The returns to academic and vocational qualifications in Britain," *Bulletin of Economic Research*, 54(3), 249–274.

Dee, T. S. (2004): "Are there civic returns to education?," *Journal of Public Economics*, 88(9), 1697–1720.

Dickson, M. (2013): "The causal effect of education on wages revisited," *Oxford Bulletin of Economics and Statistics*, 75(4), 477–498.

Dickson, M., and C. Harmon (2011): "Economic returns to education: what we know, what we don't know, and where we are going—Some brief pointers," *Economics of Education Review*, 30(6), 1118–1122.

Eren, O. (2009): "Does membership pay off for covered wokers? A distributional analysis of the free-rider problem," *Industrial and Labor Relations Review*, 62(3), 367–380.

Evans, W. N., and E. Montgomery (1994): "Education and health: Where there's smoke there's an instrument," NBER Working Paper 4949, National Bureau of Economic Research.

Fang, H., K. N. Eggleston, J. A. Rizzo, S. Rozelle, and R. J. Zeckhauser (2012): "The returns to education in China: Evidence from the 1986 compulsory education law," NBER Working Paper 18189, National Bureau of Economic Research.

Fasih, T., G. Kingdon, H. A. Patrinos, C. Sakellariou, and M. Soderbom (2012): "Heterogeneous returns to education in the labor market," Policy Research Working Paper 6170, The World Bank.

Firpo, S., N. M. Fortin, and T. Lemieux (2009): "Unconditional quantile regressions," *Econometrica*, 77(3), 953–973.

Hanushek, E. A., L. Woessmann, and L. Zhang (2011): "General ducation, vocational education, and labor-market outcomes over the life-cycle," NBER Working Paper 17504, National Bureau of Economic Research.

Harmon, C., H. Oosterbeek, and I. Walker (2003): "The returns to education: Microeconomics," *Journal of Economic Surveys*, 17(2), 115–156.

Harmon, C., and I. Walker (1999): "The marginal and average returns to schooling in the UK," *European Economic Review*, 43(4), 879–887.

——— (2000): "The returns to the quantity and quality of education: Evidence for men in England and Wales," *Economica*, 67(265), 19–35.

Hartog, J., P. T. Pereira, and J. A. Vieira (2001): "Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators," *Applied Economics*, 33(8), 1021–1037.

Heckman, J., and S. Polachek (1974): "Empirical evidence on the functional form of the earnings-schooling relationship," *Journal of the American Statistical Association*, 69(346), 350–354.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond," vol. 1 of *Handbook of the Economics of Education*, chap. 7, pp. 307–458. Elsevier.

Henderson, D. J., S. W. Polachek, and L. Wang (2011): "Heterogeneity in schooling rates of return," *Economics of Education Review*, 30(6), 1202–1214.

Hungerford, T., and G. Solon (1987): "Sheepskin Effects in the Returns to Education," *The Review of Economics and Statistics*, 69(1), 175–77.

Imbens, G. W. (2010): "Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, 48(2), 399–423.

Imbens, G. W., and J. D. Angrist (1994): "Identification and estimation of local average treatment effects," *Econometrica*, 62(2), 467–75.

Koenker, R., and G. Bassett (1978): "Regression quantiles," *Econometrica*, 46(1), 33–50.

Koop, G., and J. L. Tobias (2004): "Learning about heterogeneity in returns to schooling," *Journal of Applied Econometrics*, 19(7), 827–849.

Lamarche, C. (2011): "Measuring the incentives to learn in Colombia using new quantile regression approaches," *Journal of Development Economics*, 96(2), 278–288.

Lemieux, T. (2008): "The changing nature of wage inequality," *Journal of Population Economics*, 21(1), 21–48.

Martins, P. S., and P. T. Pereira (2004): "Does education reduce wage inequality? Quantile regression evidence from 16 countries," *Labour Economics*, 11(3), 355–371.

Maynard, A., and J. Qiu (2009): "Public insurance and private savings: Who is affected and by how much?," *Journal of Applied Econometrics*, 24(2), 282–308.

Mincer, J. A. (1974): *Schooling, Experience, and Earnings*, NBER Books. National Bureau of Economic Research.

Mwabu, G., and T. P. Schultz (1996): "Education returns across quantiles of the wage function: Alternative explanations for returns to education by race in South Africa," *American Economic Review*, 86(2), 335–39.

Park, J. H. (1994): "Returns to schooling: A peculiar deviation from linearity," Working Paper 714, Princeton University, Department of Economics, Industrial Relations Section.

Pischke, J.-S. (2007): "The impact of length of the school year on student performance and earnings: Evidence from the German short school years," *The Economic Journal*, 117(523), 1216–1242.

Powell, J. L. (1983): "The asymptotic normality of two-stage least absolute deviations estimators," *Econometrica*, 51(5), 1569–1575.

Rosenbaum, J. E., and J. Rosenbaum (2013): "Beyond BA blinders: Lessons from occupational colleges and certificate programs for nontraditional students," *The Journal of Economic Perspectives*, 27(2), 153–172.

Rupietta, C., and U. Backes-Gellner (2012): "High quality workplace training and innovation in highly developed countries," Economics of Education Working Paper Series 0074, University of Zurich, Institute for Strategy and Business Economics (ISU).

Saniter, N. (2012): "Estimating heterogeneous returns to education in Germany via conditional heteroskedasticity," IZA Discussion Papers 6813, Institute for the Study of Labor (IZA).

Staiger, D., and J. H. Stock (1997): "Instrumental variables regression with weak instruments," *Econometrica*, 65(3), 557–586.

Trostel, P., I. Walker, and P. Woolley (2002): "Estimates of the economic return to schooling for 28 countries," *Labour Economics*, 9(1), 1–16.

Tuor, S. N., and U. Backes-Gellner (2010): "Risk-return trade-offs to different educational paths: vocational, academic and mixed," *International Journal of Manpower*, 31(5), 495–519.

Vandenbussche, J., P. Aghion, and C. Meghir (2006): "Growth, distance to frontier and composition of human capital," *Journal of Economic Growth*, 11(2), 97–127.

Wang, L. (2013): "How does education affect the earnings distribution in urban China?," *Oxford Bulletin of Economics and Statistics*, 75(3), 435–454.

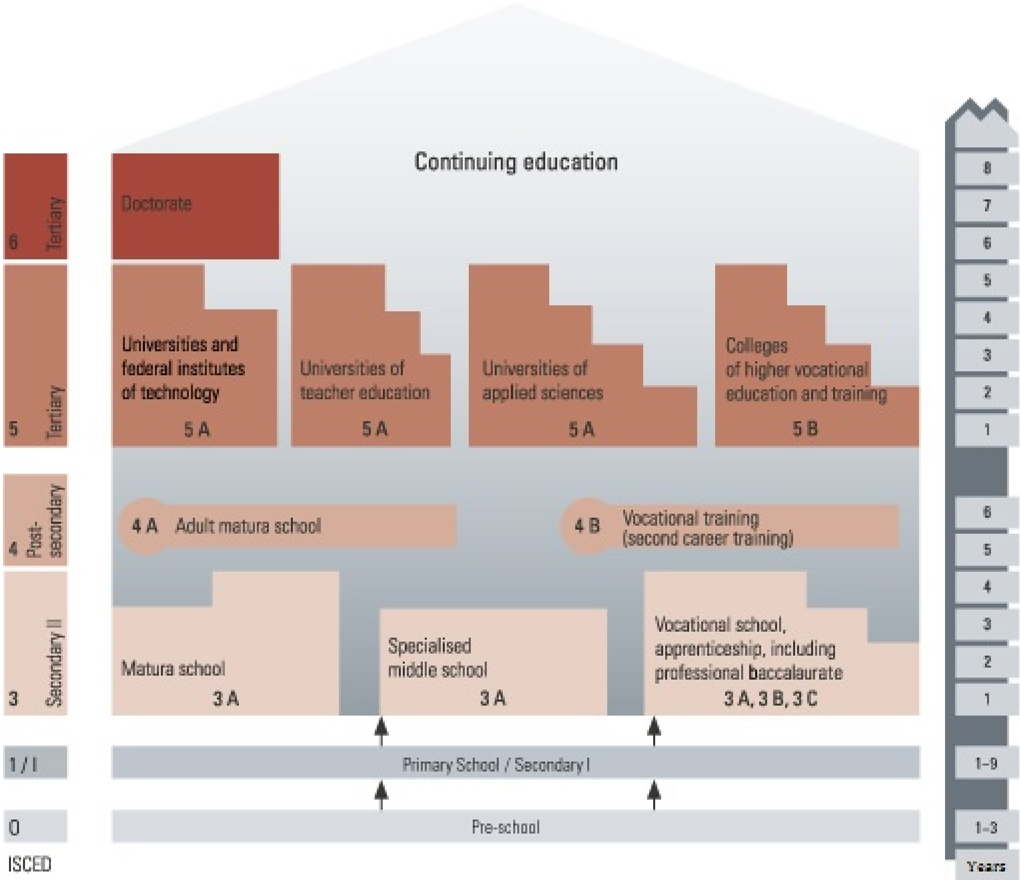
Weber, B., and S. Wolter (1999): "Wages and human capital: Evidence from Switzerland," *ETLA: The Research Institute of the Finnish Economy*.

Wehby, G. L., J. C. Murray, E. E. Castilla, J. S. Lopez-Camelo, and R. L. Ohsfeldt (2009): "Quantile effects of prenatal care utilization on birth weight in Argentina," *Health Economics*, 18(11), 1307–1321.

APPENDIX

A The Swiss Education System

Figure A.1: The Swiss Education System



B Analytic Sample

Table B.1: Sample Construction

Initial sample (SLFS 2000–2009)	160,925
Males	74,871
Fully employed	47,347
Age between 18 and 60	44,670
Not in education or gap year	42,612
Wage not missing	35,095
99 percent of wage distribution	34,744
Analytic sample	34,744

Notes: Swiss Labor Force Survey, Authors' calculations.

C Summary of the Reform of 1970

Table C.1: Compulsory Education Expansion

Canton	Entry Age (1)	Reform (2)	Year (3)	Before (4)	After (5)	First Cohort (6)
Zürich	6	Yes	1977	8	9	1971
Bern	6	No		9	9	
Luzern	6	Yes	1985	8	9	1979
Uri	7	Yes	1977	7	9	1970
Schwyz	7	Yes	1992	7	9	1985
Obwalden	7	Yes	1992	7	9	1985
Nidwalden	6	Yes	1992	7	9	1986
Glarus	6	Yes	1983	8	9	1977
Zug	7	Yes	1990	8	9	1983
Fribourg	7	No		9	9	
Solothurn	7	Yes	1970	8	9	1963
Basel-Stadt	6	No		9	9	
Basel-Land	6	Yes	1980	8	9	1974
Schaffausen	6	Yes	1982	8	9	1976
Appenzell A.	6	Yes	1981	8	9	1975
Appenzell I.	6	Yes	1984	7	9	1978
St. Gallen	6	Yes	1983	8	9	1977
Graubünden	7	No		9	9	
Aargau	7	Yes	1982	8	9	1975
Thurgau	6	Yes	1980	8	9	1974
Ticino	6	No		9	9	
Vaud	7	No		9	9	
Valais	7	Yes	1987	8	9	1980
Neuchâtel	6	No		9	9	
Genève	6	No		9	9	
Jura	6	No		9	9	

Notes: Authors' research and calculations.

Table C.2: CHANGE IN SCHOOL-YEAR START

Canton	Entry Age (1)	Reform (2)	Year (3)	Before (4)	After (5)	First Cohort (6)
Zürich	6	Yes	1989	Spring	Fall	1974
Bern	6	Yes	1989	Spring	Fall	1974
Luzern	6	No		Fall	Fall	
Uri	7	No		Fall	Fall	
Schwyz	7	Yes	1989	Spring	Fall	1975
Obwalden	7	No		Fall	Fall	
Nidwalden	6	No		Fall	Fall	
Glarus	6	Yes	1989	Spring	Fall	1975
Zug	7	Yes	1973	Spring	Fall	1958
Fribourg	7	No		Fall	Fall	
Solothurn	7	Yes	1989	Spring	Fall	1973
Basel-Stadt	6	Yes	1989	Spring	Fall	1974
Basel-Land	6	Yes	1989	Spring	Fall	1975
Schaffausen	6	Yes	1989	Spring	Fall	1975
Appenzell A.	6	Yes	1989	Spring	Fall	1975
Appenzell I.	6	Yes	1989	Spring	Fall	1976
St. Gallen	6	Yes	1989	Spring	Fall	1975
Graubünden	7	No		Fall	Fall	
Aargau	7	Yes	1989	Spring	Fall	1974
Thurgau	6	Yes	1989	Spring	Fall	1974
Ticino	6	No		Fall	Fall	
Vaud	7	Yes	1973	Spring	Fall	1957
Valais	7	No		Fall	Fall	
Neuchâtel	6	Yes	1973	Spring	Fall	1958
Genève	6	No		Fall	Fall	
Jura	6	Yes	1989	Spring	Fall	1974

Notes: Authors' research and calculations.

D Alternative IV and Over-identified Models

Table D.1: Returns to Education, TSLS Estimates

Variables	First Stage (1)	Second Stage (2)	First Stage (3)	Second Stage (4)	First Stage (5)	Second Stage (6)
Years of education		0.104**		0.101**		0.102**
		(0.026)		(0.017)		(0.014)
Age	0.166**	0.047**	0.185**	0.048**	0.179**	0.048**
	(0.011)	(0.004)	(0.011)	(0.003)	(0.011)	(0.003)
$Age^2/100$	-0.201**	-0.045**	-0.221**	-0.045**	-0.212**	-0.045**
	(0.013)	(0.005)	(0.014)	(0.004)	(0.014)	(0.003)
Constant	10.348**	8.669**	9.876**	8.703**	9.978**	8.696**
	(0.223)	(0.288)	(0.240)	(0.188)	(0.241)	(0.159)
IV_1–Education expansion			0.302**		0.620**	
			(0.055)		(0.083)	
IV_2–Shift in school start	0.221**		0.178**		0.267**	
	(0.043)		(0.044)		(0.047)	
$IV_1 \cdot IV_2$					-0.539**	
					(0.104)	
Year fixed effects	YES	YES	YES	YES	YES	YES
R^2	0.018	0.233	0.019	0.243	0.019	0.241
N	34,744	34,744	34,744	34,744	34,744	34,744
Test for excluded instruments						
F-statistic	26.69**		28.37**		26.53**	
Under-identification test						
Kleibergen-Paap LM-statistic		26.58**		55.91**		77.95**
Over-identification Test						
Hansen J-statistic (p-value)				0.860		0.982

Notes: ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Robust standard errors are in parentheses. In odd columns the dependent variable is years of education, in even columns the dependent variable is the natural logarithm of annual wage.
Swiss Labor Force Survey, Authors' calculations.

Table D.2: Over-identified IVQR

Variables	25^{th} Percentile	Median	75^{th} Percentile	25^{th} Percentile	Median	75^{th} Percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Years of education	0.178**	0.090**	0.082**	0.164**	0.085**	0.092**
	(0.032)	(0.023)	(0.025)	(0.025)	(0.019)	(0.021)
Age	0.020**	0.047**	0.056**	0.023**	0.047**	0.055**
	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)
$Age^2/100$	-0.014*	-0.045**	-0.053**	-0.017**	-0.045**	-0.051**
	(0.007)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)
Constant	8.143**	8.910**	8.979**	8.124**	8.972**	8.882**
	(0.338)	(0.250)	(0.270)	(0.281)	(0.207)	(0.228)
Year fixed effects	YES	YES	YES	YES	YES	YES
IV_1–Education expansion	X	X	X	X	X	X
IV_2–Shift in school start	X	X	X	X	X	X
$IV_1 \cdot IV_2$				X	X	X
N	34,744	34,744	34,744	34,744	34,744	34,744

Notes: ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Robust standard errors are in parentheses. The dependent variable is the natural logarithm of annual wage. Swiss Labor Force Survey, Authors' calculations.